

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-339703

(43)Date of publication of application : 07.12.2001

(51)Int.Cl.

H04N 7/15

G02B 7/08

G03B 15/00

H04N 5/232

H04R 3/00

(21)Application number : 2000-157354

(71)Applicant : NEC CORP

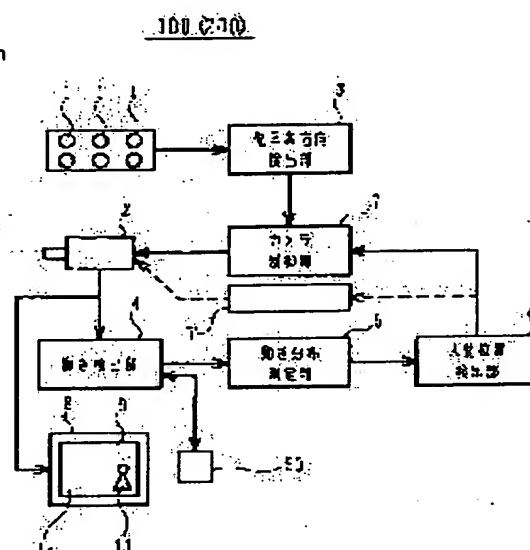
(22)Date of filing : 26.05.2000

(72)Inventor : SATODA KOZO

(54) VIDEO CONFERENCE SYSTEM, CONTROL APPARATUS OF CAMERA IN VIDEO CONFERENCE SYSTEM AND CONTROL METHOD OF CAMERA**(57)Abstract:**

PROBLEM TO BE SOLVED: To solve problems such that there is a large error when a speaker is detected only by his voice, and there is a large amount of computation because of pursuing the contour in the image by the technique suggested as a method for correcting with image after having been detected by the voice roughly in a video conference terminal.

SOLUTION: The system comprises a detector part 3 for detecting the direction of a speaker using phase difference in voice signal supplying to multi- microphones 1 and a control part 7 by which the camera 2 is pointed in the detected direction. A moving part is detected by using frame difference at motion detecting part 4 from video images picked up with the camera pointing in the direction of speaker's voice and the distribution is measured at the motion distribution measuring part 5 and then, the personal position is detected by the motion distribution in the personal position detecting part 6. Even if the speaker is not in the center position of taken image, an appropriate speaker's image enables to take by moving the camera with the camera control part, such as a detected person is in the center of the camera.

**LEGAL STATUS**

[Date of request for examination] 12.02.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration] withdrawal

[Date of final disposal for application] 26.01.2004

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

当該撮像面映出に於ける当該発着者の位置情報から、当該撮像面映出に於ける当該発着者の大きさを調整する為に、第2の撮像手段制御部が、当該撮像手段のズーム機構を調整する第7'の工程、とから構成されている事を特徴とするテレビ会議システムに於ける撮像手段の制御方法をコンピュータに実行させる為のプログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】
【発明の属する技術分野】 本発明はテレビ会議システム及びテレビ会議システムに於けるカメラの制御装置並びにカメラの制御方法に関するものであり、特に詳しくは、会議の発着者が適切に撮影されるようにカメラの撮影方向および画面を制御する制御装置に関するものである。

【0002】

【従来の技術】 従来から、テレビ会議端末において複数のカメラを用意し、テレビ会議に参加している複数の参加者の中から発着者を自動的に選択し、その発着者にカメラを向けるための制御装置が開発されてきた。例えば、特開平5-122689号公報で公開されている方法では、各参加者に対応するようにカメラを配置しておき、音レベルが最大のカメラの方向にカメラを向けることで発着者をカメラで撮影している。

【0003】 また、特開平7-140627号公報で公開されている方法では、複数のカメラに入力される音声の位相の違いから音の聞こえてくる方向を抽出している。

【0004】 一方、カメラで撮影された画像から人物を抽出し、その方向にカメラを向ける方法も開示されている。例えば、特開平4-234284号公報では画像のフレーム間差分を用いて、水平方向の動きを分布を特定することにより、画像中の人の位置を抽出している。

【0005】 また、特開平5-268598号公報では画像から輪郭を抽出し、人の形のパターンとテンプレートを照合することにより人物を抽出している。

【0006】 また、演算に参加者を撮影するために、広角カメラを別に用意する方法が、例えば、特開平5-244687号公報などで開示されている。

【0007】 又、特開平8-29662号公報のように、音声で人物を抽出した後、画像から人物の輪郭を抽出し、カメラの方向を補正する手法が開示されている。

【0008】

【発明が解決しようとする課題】 然しながら、上記した従来の技術のように複数カメラの中から音レベルが最大のものにカメラを向けるだけでは、カメラの方向と発着者の方向が一致している場合しか使用できず、発着者の方向がカメラの方向とずれている場合には発着者を中央に撮影することができない。

【0009】 また、音声の位相差による手法は精度が大

きく、カメラの中央に発着者をとらえることが望ましいという欠点がある。

【0010】 その他、画像によりカメラ中央に発着者をとらえる方法は、画像中に人物が映っていない場合は適用できず、更には、広角カメラを人物撮影用のカメラと別に用意する手法では、テレビ会議システム全体が高価になってしまう。

【0011】 一方、音声で狙った撮像面を補正する方法として提案されている手法では、画像中の輪郭を求め、そこから人物を抽出する。しかし、輪郭を用いる方法は会議室内の様々な物体が映っているとき計算量が大きくなるという問題がある。

【0012】 また、この手法では人物を適切な画面内で撮影する手段については考慮していない。

【0013】 従って、本発明の目的は、上記した従来の技術の欠点を改良し、簡易な機構でありながら、発着者の発着する音声情報を基に、当該発着者の方向に正確に且つ正確に撮像手段であるカメラを向けると共に適切な画面内で発着者を撮像できるテレビ会議システムを提供するものである。

【0014】

【課題を解決するための手段】 本発明は、上記した目的を達成するため、以下に記載されたような技術構成を採用するものである。即ち、本発明に係る第1の態様としては、複数の集音手段、少なくとも一つの発着者検出手段、画像表示手段、及び当該集音手段から得られる発着者の音声方向情報に基づいて、当該発着者を撮像する撮像手段の撮像方向を変更する撮像制御手段とから構成されているテレビ会議システムに於ける、当該集音手段により検出された発着者の検出位置位置に向けて当該撮像制御手段が当該撮像手段の撮像方向を指向せしめる様に構成すると共に、当該撮像面映出から動きを抽出し、その動きを画像の分布を求め、当該発着者の位置情報に基づいて、当該発着者が当該撮像面の所定の部位に表示される様に、当該撮像制御手段を更に制御する様に構成したテレビ会議システムであり、又、本発明に係る第2の態様としては、複数の集音手段、発着者を撮像する撮像手段、当該集音手段からの情報に基づいて発着者の存在位置を特定する発着者位置検出手段、当該発着者の存在位置の情報に基づいて、撮像手段の指向方向を変化させる第1の撮像手段制御部、当該第1の撮像手段制御部の制御信号に応じて所定の方向に指向せしめられた当該撮像手段の撮像面を指示する画像表示手段、当該撮像面映出画像から動きを抽出する動き画像抽出手段、当該動き画像抽出手段により抽出された動き画像から動きを分布を特定する動き分布検出手段、当該動き分布検出手段の検出結果に基づいて、発着者の当該撮像面に於ける存在位置を決定する発着者存在位置決定手段、当該発着者存在位置決定手段の情報に基づいて、当該撮像手段の

指向方向を更に制御する第2の撮像手段制御部とから構成されているテレビ会議システムに於ける撮像手段の制御装置である。

【0015】 更には、本発明に係る第3の態様としては、複数の集音手段、少なくとも一つの発着者検出手段、発着者表示手段及び当該集音手段から得られる発着者の音声方向情報に基づいて、当該発着者を撮像する撮像手段の撮像方向を変更する撮像制御手段とから構成されているテレビ会議システムに於いて、発着者の音声と当該集音手段のそれぞれが検出した情報から当該発着者の存在位置を特定する第1の工程、当該第1の工程に於いて、指定された当該発着者の存在位置方向情報に基づいて、第1の撮像手段制御部が当該撮像手段を駆動させて、当該撮像手段の撮像方向を当該発着者の検出位置位置に向けて指向せしめる第2の工程、当該撮像手段が撮像する画像を画像表示装置上に表示する第3の工程、当該撮像面映出画像から、動き画像情報を抽出する第4の工程、当該抽出された動き画像情報から発着者の当該撮像手段の第6の工程、当該動き分布情報から発着者の当該撮像面映出中に於ける存在位置を決定する第6の工程、当該撮像面映出中に於ける当該発着者の位置情報から、当該撮像面映出に於ける当該発着者の大きさを調整する為に、第2の撮像手段制御部が、当該撮像手段のズーム機構を調整する第7'の工程、とから構成されているテレビ会議システムに於ける撮像手段の制御方法である。

【0016】

【発明の実施の形態】 本発明に係るテレビ会議システム及びテレビ会議システムに於けるカメラの制御装置並びにカメラの制御方法は、上記した様な技術構成を採用している。このため、複数のカメラに入力される音声の位相差あるいは音レベル等から発着者の方向を抽出する発着者方向検出部と、抽出した方向にカメラを向けるカメラ制御部を持つものであって、音声により発着者方向に向けられたカメラが撮る映像の中から動き抽出部で動きのある部分を抽出し、その分布を動き分布検出部により水平方向、垂直方向に対して検出し、水平方向、垂直方向の動きの分かれから人物位置検出部で人物位置、大きさを抽出する。発着者が画像の中央に映っていない場合でも、発着者を適切な大きさでカメラ中央に撮影できるようにカメラ制御部でカメラを移動することができ。

【0017】

【実施例】 以下に、本発明に係るテレビ会議システム及びテレビ会議システムに於けるカメラの制御装置並びにカメラの制御方法の一例の構成を図面を参照しながら詳細に説明する。

【0018】 即ち、図1は、本発明に係るテレビ会議システムの一具体例の構成の概要を示すブロックダイアグラムであって、図中、複数の集音手段1、少なくとも一つの発着者検出手段2、及び当該集音手段1から得られる発着者の音声方向情報に基づいて、当該発着者を撮

像する撮像手段2の撮像方向を変更する撮像制御手段7及び発着者の画像を表示する画像表示手段8とから構成されているテレビ会議システム100に於いて、当該集音手段1により検出された発着者の検出位置位置に向けて当該撮像制御手段7が当該撮像手段2の撮像方向を指向せしめる様に構成すると共に、当該撮像面映出から動き画像を抽出して、その動き画像の分布を求める事によって当該発着者の位置情報12上の位置を特定し、当該発着者の位置情報に基づいて、当該発着者11が画像表示領域12の所定の部位に表示される様に、当該撮像制御手段7を更に制御する様に構成したテレビ会議システム100が示されている。

【0019】 つまり、本発明に係る当該テレビ会議システム100のより具体的な構成を示すならば、本発明では、例えば、複数のマイク11に入力される音声の位相差から発着者の方向を抽出する発着者方向検出部3と、抽出した方向にカメラを向けるカメラ制御部7をもち、当該音声により発着者方向に向けられた撮像手段の一つであるカメラが撮る撮像面映出の画像情報の中から動きを抽出する動き検出部4をもち、動き検出部4から動きを抽出し、その動きを動き分布を求め、動き分布検出手段5により水平方向、垂直方向の何れか若しくは双方に対して検出し、水平方向、垂直方向の動きの分布から人物位置検出部6つまり、発着者存在位置決定手段6で人物位置、大きさを抽出する。

【0020】 本発明においては、当該発着者の画像が、当該撮像面映出中にあるかあるいはその位置を特定する場合、上記した表示手段8に当該画像を表示せず、当該動き検出手段4内に設けられた適宜の画像処理回路内に格納されたデータは当該動き検出手段4に投与された任意の記憶回路600を使用する事によって処理する事ができる。

【0021】 尤も、当該表示手段8に送致当該発着者の像を表示しておくことも可能である。

【0022】 当該発着者が当該撮像面映出領域の所定の部位、例えば中央部に映っていない場合でも、当該抽出された当該発着者の表示画面以上の位置情報に基づいて、当該発着者の大きさを適切なサイズに調整する事も含めて、当該発着者の画像をカメラ中央に来る様に撮影できるようにカメラ制御部7でカメラを移動あるいはズーム機構の調整等の制御処理を行う事ができる。

【0023】 即ち、当該発着者が、当該撮像面映出中に予っていない場合には、当該撮像手段が当該カメラ2のズーム機構を引く事によって、撮像範囲を拡大して、当該発着者が当該撮像面映出中に入る様に処理する事が出来る。

【0024】 一方、当該発着者が既に当該撮像面映出内に存在している場合でも、当該発着者の表示画面映出での大きさを変化させる必要がある場合には、同時に当該撮像手段であるカメラ2のズーム機構をズームアップする

15

$|Y(x, y, n) - Y(x, y, n-1)| < T$ の場合
D(x, y, n) = 0
 $|Y(x, y, n) - Y(x, y, n-1)| \geq T$ の場合
D(x, y, n) = 1
この式中で | | は値の絶対値をとることを示す。
[0063] Dの値は発音者が動いている箇所では1であり、動かない部分では0となる。
[0064] ここではフレーム間差分を取るために前のフレーム(n-1)との差分を使ったが、更に前のフレームを使用してもよい。

[0065] 1フレーム前の画像との差分を使用した場合、動きが小さい部分が抽出されないなどの問題があるため、前の画像を覚えておくための記憶領域が多く必要となるが、安定して動き部分を抽出することができ、[0066] 動き抽出部4で抽出した動きD(x, y, n)は動き分布判定部6に送られ、ここで水平方向および垂直方向の動き分布を決定する。

[0067] これは図3(b)の差分画像から図3(c)の動きの分布を求める処理である。
[0068] 水平方向の動きをH(x, n)、垂直方向の動きをV(y, n)とすると、

$$H(x, n) = \sum_{y=1}^Y D(x, y, n)$$

$$V(y, n) = \sum_{x=1}^X D(x, y, n)$$

である。ここで、w、hは画像の水平方向の大きさと垂直方向の大きさである。

[0069] このようにして求めた動き分布をヒストグラム3.3、3.4として求め、当該ヒストグラムは、発音者が動いている場所では大きな値を示し、動いていない箇所では0や小さな値を示す。

[0070] 発音者位置抽出部6では動き分布判定部6で求めたH(x, n)およびV(y, n)を使って発音者の位置を抽出する。

[0071] H(x, n)の分布の中で最も大きなH(x, n)を示すxを発音者の水平方向の位置MXとすると、

[0072] 動きの分布はこの最大のMXを中心に、MXから隔れるほどH(x, n)の値が小さくなるので、ある程度まで小さくなったところまでの範囲(MX1~MX2)を水平方向に人が存在している範囲であるとする。

[0073] また、垂直方向も同様に最大のMYとそこから隔れてある程度まで小さくなったところまでの範囲(MY1~MY2)を垂直方向の分布であるとする。

16

[0074] MXからカメラの水平方向の移動量(パツ)が計算できる。MX2は発音者の大きさをあらわしているため、発音者を画面内に適当な大きさで撮影するための画角の補正量が計算できる。

[0076] MY1、MY2からカメラの垂直方向の移動量(デルタ)が計算できる。ここで計算したデルタ、画角で人物を撮影するようにカメラ制御部7に要求することで、カメラ2が発音者を適切な大きさと画面の中央に撮影することができ、

[0076] 以上の具体例ではフレーム間差分で位置信号Y(x, y, n)しか用いなかったが、背景と前景や皮膚の輝度成分が収まっていることもあるので、色信号を用いると差分による動き抽出の精度が上がる。

[0077] また、動き抽出部4以後の処理は画像中に人物映像が映っていない場合は、カメラの方向と画角を補正することができないので、カメラ制御部7にカメラの移動を要求すると同時に、画角を十分広くするようにカメラを制御するようにすると、確実に発音者の映像を画面内におさめることができる。この画角は音声による抽出の角度調整の値に合致するとよい。

[0078] 次に、本発明に於ける当該テレビ会議システムに於ける撮像手段の制御装置200の操作方法の一例を、第5の具体例として、図4を参照しながら説明する。

[0079] 即ち、本発明に係る当該テレビ会議システムに於ける撮像手段の制御装置の操作手順は、図4に示す様に、複数回の発音手段1、少なくとも一つの発音者撮像手段2、画像表示手段8及び少なくとも当該発音手段1から得られる発音者の音声方向情報に基づいて、当該発音者を撮像する発音者の音声方向情報に基づいて、当該発音者撮像部7とから構成されているテレビ会議システム100に於いて、発音者の音声当該各発音手段1のそれぞれが集音した情報から当該発音者の存在位置を決定する第1の工程、当該第1の工程に於いて、決定された当該発音者の存在位置方向情報に基づいて、第1の撮像手段2の撮像方向を当該発音者の指定存在位置に向けて指向せしめる第2の工程、当該撮像手段2が撮像する画像を撮像画像として記憶する第3の工程、当該記憶された撮像画像から、動き画像情報3.1を抽出する第4の工程、当該抽出された動き画像情報3.1から動き分布3.2を算出する第5の工程、当該動き分布情報3.2から発音者の当該撮像画像中に於ける存在位置3.6を決定する第6の工程、当該表示手段8中に於ける当該発音者の位置情報から、第2の撮像手段制御部7が、当該撮像手段2を移動させて、当該発音者の画像を当該撮像画像中の所定の位置に移動させる第7の工程、とから構成されているテレビ会議システムに於ける撮像手段の制御方法である。

[0080] 本発明に於ける当該テレビ会議システムに

17

於ける撮像手段の制御方法に於いては、当該第1の撮像手段制御部7と当該第2の撮像手段制御部7とは同一の制御部であってもよく、又、当該第1の工程に於ける発音者の音声方向情報に基づき、当該撮像手段2の撮像方向の位置は、当該撮像手段1又は13のそれぞれに入力される音声の位相差を用いた音声レベルに基づいて決定されるものである。

[0081] 又、本発明に於ける当該動き画像情報を抽出する第4の工程は、撮像された異なる画像フレーム間から差分画像を作成するものである事が望ましい。

[0082] 一方、本発明に於ける当該差分画像3.1は、逐次撮像される撮像画像をフレーム毎に適宜の記憶手段6.0に記憶せしめ、当該記憶手段6.0に記憶された表示画像フレームから選択された異なる表示画像フレーム間、若しくは当該記憶手段に記憶された一つの表示画像フレームと現在得られている撮像画像に於ける画像フレームとの間の差分値を求めるものである。

[0083] 本発明に於ける当該抽出された動き画像情報から動き分布を算出する第5の工程は、当該抽出された動き画像情報から、少なくとも水平方向若しくは垂直方向の何れか一方について当該画像情報に関するヒストグラムを形成し、当該ヒストグラムを予め定められた基準ヒストグラムと比較して、当該発音者の当該撮像画像上に於ける存在位置を識別するものである事が望ましい。

[0084] 本具体例に於て、当該撮像画像は、必要に応じて当該画像表示手段8の画像領域1.2に表示する様にする事も可能である。

[0085] 更に、本発明に於いては、当該動き分布情報から発音者の当該撮像画像中に於ける存在位置を決定する第6の工程に於て、当該発音者の存在位置が識別出来なかった場合には、当該撮像手段2により現在撮像している領域を拡大して、当該第4乃至第6の各工程を繰り返す操作が実行される必要がある。

[0086] 具体的に、例えば、当該撮像手段2により現在撮像している領域を拡大する工程は、当該撮像手段に於けるズームを引く操作を実行するものである。

[0087] 尚、本発明に於いては、発音者の画像上での大きさを、大きくしたり小さくしたりする等任意に変化させるズーム操作を任意に実行することにより可能である。

[0088] 又、本発明に係る第6の具体例としては、撮像手段の発音手段、少なくとも一つの発音者撮像手段、発音者表示手段及び当該発音者撮像手段から得られる発音者の音声方向情報に基づいて、当該発音者を撮像する撮像手段の撮像方向を変更する撮像手段制御部7が、当該撮像手段のテレビ会議システムに於いて、発音者の音声当該各発音手段のそれぞれが集音した情報から当該発音者の存在位置を決定する第1の工程、当該第1の工程に於いて、決定された当該発音者の存在位置方向情報に基づいて、

(10)

18

て、第1の撮像手段制御部が当該撮像手段を移動させて、当該撮像手段の撮像方向を当該発音者の指定存在位置に向けて指向せしめる第2の工程、当該撮像手段が撮像する画像を画像表示手段8.1に表示する第3の工程、当該撮像手段制御部から、動き画像情報を抽出する第4の工程、当該抽出された動き画像情報から動き分布を算出する第5の工程、当該動き分布情報から発音者の当該撮像画像中に於ける存在位置を決定する第6の工程、当該撮像手段制御部7に於ける当該発音者の位置情報から、当該撮像手段制御部7が、当該撮像手段2を移動させて、当該発音者の画像を当該撮像画像中の所定の位置に移動させる第7の工程、とから構成されているテレビ会議システムに於ける撮像手段の制御方法である。次に、本発明に於ける当該テレビ会議システムに於ける撮像手段の制御装置200の操作方法の一例を、第7の具体例として、図6に示すフローチャートを用いながら説明する。

[0089] 即ち、スタート後、ステップ(S1)に於て、会議が開始されたか否かが判断され、YESであれば、ステップ(S2)に進み、上記の操作が繰り返され、YESであれば、ステップ(S4)に於て、当該発音手段1の方向に発音者が存在すると判定する前記第1の工程が実行され、ステップ(S5)で、前記した第1の撮像手段制御部7を起動させて、当該撮像手段2の撮像方向を当該発音者の指定存在位置に向けて指向せしめる第2の工程が実行される。

[0091] 続いて、ステップ(S6)に於て、当該撮像手段2の撮像方向と当該発音者がいる指定位置とが一致したか否かが判断され、NOであれば、ステップ(S7)に進み、当該撮像手段2が撮像する画像を撮像画像として記憶する第3の工程が実行された後、ステップ(S8)に進み、当該撮像画像から、動き画像情報3.1を抽出する第4の工程が実行される。

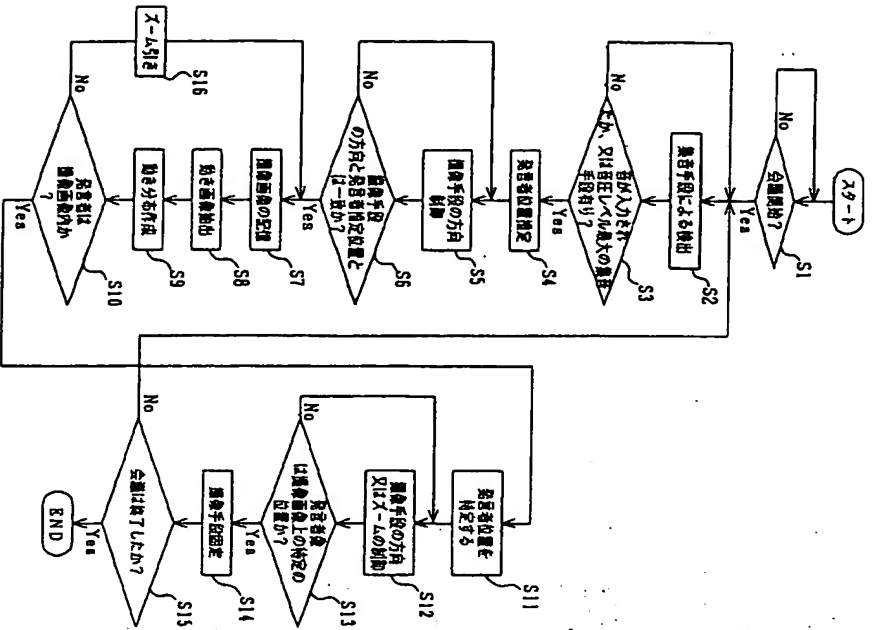
[0092] その後、ステップ(S9)に於て、当該抽出された動き画像情報3.1から動き分布3.2を算出し、ヒストグラムを求める第5の工程が実行される。

[0093] 次に、ステップ(S10)に進み、当該発音者の存在位置が当該撮像画像中に有るか否かが判断され、NOであれば、ステップ(S16)に進み、当該撮像手段2にズーム操作を行わず、(この場合はズームを引く操作が繰り返される。

[0094] 当該ズーム(S10)でYESであれば、

60

【図6】



フロントページの続き

(51)Int. Cl.⁷
H04R 3/00 320

識別記号

F1
H04R 3/00

320
フーコ-1' (備考)